

**PATENT****REMARKS**

Reconsideration of the rejections set forth in the Office action dated 6/15/2005 is respectfully requested under the provisions of 37 CFR §1.111(b).

Applicant thanks the Examiner for the indication of allowability of claims 2-6, 8, 13-17 and 19.

Claims 1-20 are pending.

Claims 1, 7, 9-12, 18 and 20 stand rejected.

Claims 2-6, 8, 13-17 and 19 are objected to.

No claims were amended.

***I. Rejections under 35 USC §103***

Claims 1, 10, and 11-12 stand rejected under 35 USC §103(a) as being unpatentable over the combination of Lopresti et al. (5,748,807) and Visser (6,023,536). This rejection is respectfully traversed as a prima facie case of obviousness has not been presented.

A prima facie case of obviousness is established by one or more references that were available to the inventor and that teach a suggestion to combine or modify the reference, the combination or modification of which would appear to be sufficient to have made the claimed invention obvious to one of ordinary skill in the art.

**A. Overview**

One aspect of the claimed invention is described starting at page 41, line 13 and describes a "rewrite model" of an embodiment that uses aspects of the inventive techniques. The preceding portion of the instant application has been directed to inventive techniques for correcting optical character recognition (OCR) errors where the OCR recognition program incorporates an assist channel and a primary channel; where

**PATENT**

the assist channel provides the OCR recognition program with additional information that can be used by the OCR recognition program to correct errors or omissions of the character images (page 8, lines 21). Three embodiments for encoding and decoding the assist channel (the separation coding, block coding and convolution coding) have been disclosed prior to page 41. These techniques were described as being incorporated within the OCR decoder module 216 itself (see fig. 2, and page 7, line 18 – page 8, line 2). That is, these descriptions were of embodiments where the OCR function and disclosed use of the assist channel were combined.

The “rewrite model” is directed to embodiments where the OCR function in the decoding module is independent of the assist channel correction function. Thus, an off-the-shelf OCR program (decoding module) can be used to recognize characters in the primary channel and the decoding module output (the candidate primary channel) can be adjusted using the assist channel independently from the decoding module (page 41, lines 22-28). Notice that there is no ambiguity in the characters in the candidate primary channel. The decoding module has provided an unambiguous (although possibly incorrect) character.

One aspect of the rewrite model is an event library that identifies potential OCR error modes and provides correction rules and weightings for those error modes (page 42, lines 7-20). The rules in the event library coupled with the information in the assist channel allows this aspect of the invention to detect and repair decoding errors in the candidate set of symbol data generated by the OCR module.

The event library consists of rules that specify transformations if the secondary set of encoding data (the assist channel data) indicate that an error exists in the candidate set of symbol data. The rules in the event library are triggered on the detection of an event (that being the detection of an inconsistency between the information in the assist channel and a symbol in the candidate set of symbol data. Once an event is triggered, a rule from the event library is selected responsive to the inconsistent character and the corresponding information in the assist channel. The claimed invention does not require that the symbol data be recognized words or words in a dictionary. The claimed invention, by way of the assist channel, detects that an error exists in the decoding of the

**PATENT**

scanned representation and fixes that error such that the decoded symbol data is an accurate reconstruction of the symbols from the scanned representation. Misspelled words in the decoded symbol data will be rewritten with the same misspelling they had in the original scanned document.

This aspect of the invention is claimed by previously presented **claims 1, 10, and 11**. Claim 1 is:

1. A method for decoding image data for a hardcopy document, comprising:

recording a scanned representation of the hardcopy document that includes a primary set of symbol data and a secondary set of encoding data; the primary set of symbol data providing a first channel of human readable information rendered on the hardcopy document; the secondary set of encoding data providing a second channel of machine readable information rendered on the hardcopy document;

receiving a decoded form of the scanned representation of the hardcopy document from a decoding module to define a candidate set of symbol data; and

rewriting, independent of the decoding module, the candidate set of symbol data using an event library and the secondary set of encoding data; the event library identifying likely failures encountered when the scanned representation of the hardcopy document is decoded; the event library comprising a rule that represents a transformation.

**B. Cited Art**

**Lopresti:** The Lopresti reference discloses techniques for adding additional information to a page that will be scanned. When the page is scanned the OCR conversion can use the additional information (for example, to detect and correct errors in the conversion). Lopresti does not teach rewriting the candidate set of symbol data using an event library (that comprises a rule that represents a transformation) and the secondary

**PATENT**

set of encoding data. Nor does Lopresti teach the event library identifying likely failures encountered when the scanned representation of the hardcopy document is decoded. Furthermore, nothing in Lopresti teaches a suggestion to make such a combination.

**Visser:** The Visser reference discloses a spelling correction process that teaches techniques for detecting and correcting misspellings of a string by using a dictionary and analyzing the structure of the words in the string (morphological analysis). Visser reads the string character-by-character and compares the read string to an error pattern. If the string matches the error pattern, a spelling error is assumed and the just-read character is replaced by an alternative character. (Column 6, lines 47-55; column 7, lines 50-61). Each time a character is read, the state of the analysis is saved (column 6, lines 55-57; column 8, lines 53-55). The saved states are checked against the dictionary for recognized words and presented (column 8, lines 60-67).

Visser is a spelling correction system. Thus, if an incorrectly spelled word (or a string that was not in the dictionary) was included in a document that was scanned, and Visser was applied to the OCR representation of the document, the misspelled word would be corrected and the OCR representation would not be an accurate reconstruction of the original document.

Visser does not teach nor suggest a secondary set of encoding data that provides a second channel of machine readable information rendered on a hardcopy document. Further, Visser does not teach nor suggest rewriting the candidate set of symbol data using an event library and the secondary set of encoding data at least because the claimed invention requires the secondary set of encoding data to be recorded on a hardcopy document.

As discussed above, the event library (that comprises a rule that represents a transformation) and the secondary set of encoding data is used to detect and correct an error made by the decoding module. The rules in the event library are triggered on the detection of an event (that being the detection of an inconsistency between the information in the assist channel and a symbol in the candidate set of symbol data). Once an event is triggered, a rule from the event library is selected responsive to the

**PATENT**

inconsistent character and the corresponding information in the assist channel. This technique allows the rewritten candidate set of symbol data to be an accurate reconstruction of the symbols that were originally scanned.

Furthermore, nothing in Lopresti and Visser would suggest the combination to one skilled in the art because the combination would not provide an accurate reconstruction of the originally scanned symbol data.. Thus, **currently amended claims 1, 10 and 11 are patentable.**

Previously presented **claim 12** depends on and further limits claim 11 that is patentable. Thus, claim 12 is patentable.

Claims 7, 9, 18 and 20 stand rejected under 35 USC §103(a) as being unpatentable over the combination of Lopresti et al. (5,748,807) and Visser (6,023,536) and Kopec (5,594,809). This rejection is respectfully traversed as a prima facie case of obviousness has not been presented.

Previously presented claims 1 and 11 are patentable.

Claims 7, 9, 18 and 20 depend from and further limit their respective independent claim. Thus **claims 7, 9, 18 and 20 are also patentable.**

In summary, previously presented claims 1, 10 and 11 are patentable for the reasons above. Original claims 2-9 depend on and further limit (either directly or through an intervening claim) claim 1. **Thus claims 2-9 are patentable.** Claims 12-20 depend on and further limit (directly or through an intervening claim) currently amended claim 11 that is patentable. Thus, **claims 12-20 are patentable.**

Since all rejections, objections and requirements contained in the outstanding official action have been fully answered or traversed and shown to be inapplicable to the present claims, it is respectfully submitted that reconsideration is now in order under the provisions of 37 CFR §1.111(b) and such reconsideration is respectfully requested. Upon reconsideration, it is also respectfully submitted that this application is in condition for allowance and such action is therefore respectfully requested.

## PATENT

No additional fee is believed to be required for this amendment. However, the undersigned Xerox Corporation attorney hereby authorizes the charging of any necessary fees, other than the issue fee, to Xerox Corporation Deposit Account No. 24-0025. This also constitutes a request for any needed extension of time and authorization to charge all fees therefor to Xerox Corporation Deposit Account No. 24-0025.

Should any additional issues remain, or if I can be of any additional assistance, please do not hesitate to contact me at (650) 812-4259.

Respectfully submitted,



DANIEL B. CURTIS  
Attorney for Applicants  
Reg. No. 39,159  
(650) 812-4259  
dbcurtis@parc.com